The Use of Computer Modeling Packages to Illustrate Uncertainty in Risk Assessments

Marc L. Rigas[1] and David M. Hassenzahl[2]

1. U.S. EPA, Office of Research and Development, NERL, PO Box 93478, Las Vegas, NV 89193-3478

2. Department of Environmental Studies, University of Nevada, Las Vegas, Las Vegas, NV 89154

OBJECTIVES

The goals of this study were to evaluate how scientists and lay people:

1) Understand the advantages and shortcomings of point estimates of exposure and risk.

2) Understand the advantages and shortcomings of Monte Carlo (MC) simulation in exposure and risk analysis.

3) Explore how different software interfaces produce Monte Carlo (MC) simulation outputs that is meaningful.

METHODS

Eleven individuals participated in this pilot level study. Almost all were employees of the U.S. Environmental Protection Agency's Office of Research and Development, Three were either Fellowship students working at EPA or graduate students in the Environmental Studies Program at the University of Nevada, Las Vegas. Criteria for selecting individuals was self-reported understanding of basic statistics in analyzing scientific data.

After providing informed consent, all participants were given the same case study to read (Fig 1). The case study was based on research by Finkel *et al.* (1995) as presented by Kammen and Hassenzahl (1999). Study sessions proceeded as follows:

- Initially, participants were given average values for food consumption and chemical residues for both apples and peanut butter. There is uncertainty associated with the cancer potencies, and average values were given for these as well. Participants were asked to perform the calculations and come up with a policy recommendation on the regulation of Alar when confronted by critics arguing that natural toxins in foods such as peanut butter pose at least as much risk.
- Participants were given a brief explanation of MC simulation and were assigned to one of three MC software packages: *@Risk* (Palisade Corp.; Newfield, NY); *Crystal Ball* (Decisioneering; Denver, CO); or *Modelmaker 4.0* (modelkinetix.com, Cambridge, UK). Four participants were assigned to each software package (only three used *Crystal Ball).*
- After standardized software instruction, participants executed a pre-setup MC simulation of the risk problem for 5000 runs.
- Participants were instructed to manipulate the output of the software package to further evaluate the relative risks of Alar and aflatoxin. They were asked:
  1. If results from MC simulation differed significantly from their original results obtained using average values.
  2. What their policy recommendations were regarding Alar regulation and how they would defend against critics asked after participants used average values).
  3. If they were able to produce meaningful output from the software and to print an example of such output.
- Finally, participants were asked to numerically rate the ease of understanding the concept of MC simulation and the ease of using their assigned software package on a scale of 1 - 5 (1 most difficult, 5 easiest).

FIGURE 1. The Comparative Risk Analysis case study presented to study participants. Parameter averages and complete distributions were based on Finkel (1995) as reported in Kammen and Hassenzahl (1999).

Case Study: Comparative Risk of Alar and aflatoxin exposure

You are a government policy maker tasked with deciding whether or not to regulate the infamous chemical, Alar, a growth regulator often used on apples processed for use in apple juice. As is common, critics of proposed regulation try to compare the risks of ingesting the chemicals with other "natural" risks to which people are exposed every day.

- Critics have charged that aflatoxin, a naturally occuring cancer-causing contaminant in peanut butter, causes more of a health risk than does Alar. Consequently, they ask "why regulate Alar."
- Proponents of the regulation claim that the decision to regulate alar should be independent of other risks. Alar clearly has adverse health effects and should, therefore, be regulated (and ideally banned).

You have asked a team of your best scientists to study this issue, and they compile data and present you with the following information:

UDMH, a decay product of Alar, is associated with toxic effects in laboratory animals.

A linear low-dose model is appropriate for predicting the cancer risk associated with either of these chemicals, the added risk (R) for eating contaminated food-products is given by:

$$R = \frac{(A \cdot C \cdot \beta)}{W_B} \qquad (1)$$

where A is the amount of a food eaten per day, C is the concentration of the chemical on the food, and β is the cancer potency of the chemical. W_B is the bodyweight of the individual. For example, if R=0.02, or 2/100 this means that the chemical exposure is expected to cause two additional cancers in 100 people.

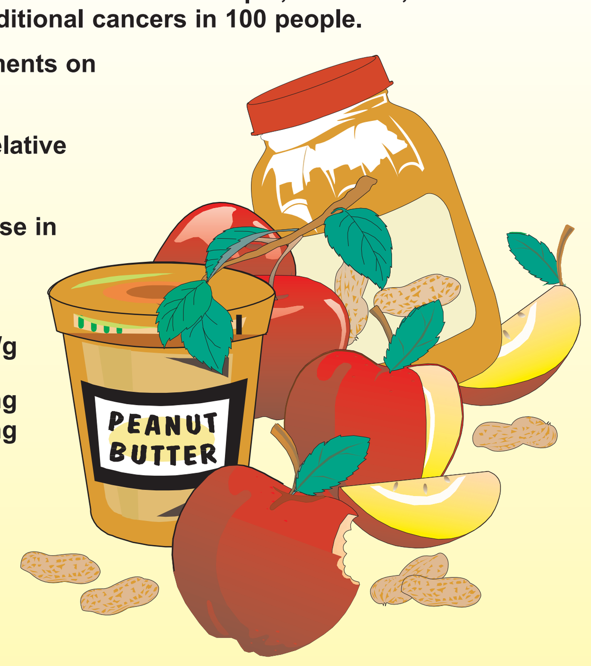
The β's for both aflatoxin and UDMH result from similar experiments on laboratory animals.

To compare risks, one can compute a ratio of R's, that is, the relative risk of one chemical to another.

Your team also found out the following information for you to use in making your judgment:

Average peanut butter consumption in the U.S. :	11.38 g/day
Average apple juice consumption in the U.S.:	136.84 g/day
Average conc of aflatoxin in peanut butter:	2.82×10^{-3} mg/g
Average conc of UDMH in apple juice:	0.01375 mg/g
Cancer potency of aflatoxin:	17.5 kg-day/mg
UDMH potency:	0.49 kg-day/mg

(Assume that neither opponents nor proponents of regulation contest the data provided here).



RESULTS

Preliminary Calculations

All but one of the subjects calculated the average risk values correctly. Average values indicated that the cancer risk of Alar in apple juice was 0.046 and the risk of aflatoxin in peanut butter was 0.024. Both of these values are well above standard regulatory thresholds, however this information was not presented to participants.

All participants except one said that they would regulate Alar based on the average risk information. One participant (ID: CB02) reported that there was insufficient information based on the averages to make a regulatory decision.

Only two of the participants compared the Alar risk to Peanut Butter in their justification for regulating Alar.

Software Use

Regulation decisions based on the MC simulations were less consistent than they were when participants were asked to look only at the average values. 3 of the 4 people who used @Risk maintained their initial decision to regulate alar, but 2 said they were presented with more clear information with which to make that decision. 1 person chose not to regulate Alar based on the new simulation results.

With *Crystal Ball,* 2 people elected to still regulate Alar, although one said that PB risk appeared more "widespread". One person (ID: CB02) noted that after seeing the percentiles of risk, they would not choose to regulate Alar. Figure 2 contains the output that led to that decision.

Modelmaker users found the output generally least useful. Three participants reported that the output didn't tell them anything new. One of those said that the output confused them so this person could no longer make a regulation decision. One person said that the simulation results made Alar risk seem less significant and chose not to regulate.

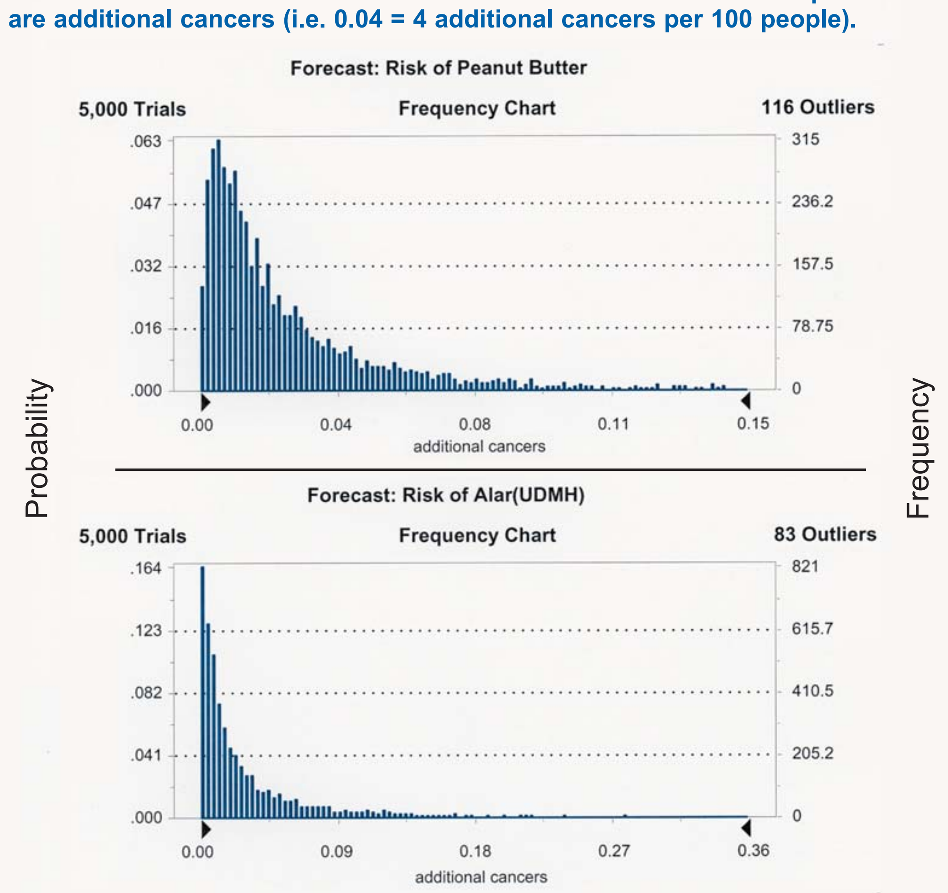
The results of participants quantitative evaluations are presented in Table 1. While the number of participants is too small to detect significant differences between each of the software packages, the concept of MC simulation was generally reported to be easy to understand by all study participants (mean = 4.58, =0.51).

Modelmaker received the highest quantitative scores. However, Figure 3 indicates that reported ease of use does not necessarily coincide with production of useful output.

	@Risk (n=4)	*Crystal Ball* (n=3)	*Modelmaker* (n=4)
Concept of MC Simulation	4.50 ± 0.50	4.67 ± 0.47	4.75 ± 0.43
Ease of understanding results	3.00 ± 0.71	4.00 ± 0.82	4.25 ± 1.30
Ease of manipulating and viewing results	3.75 ± 0.43	3.67 ± 0.47	4.25 ± 1.30

TABLE 1. Average numeric score ratings for each software package on ease of use (Score of 1 indicates difficult and a score of 5 is easiest).

FIGURE 2. Output from Crystal Ball produced by a participant (ID: CB 03). This output consists of frequency histrogram plots for a MC simulation of 5000 risk calculations for both Alar risk and aflatoxin risk. Risk vallues presented are additional cancers (i.e. 0.04 = 4 additional cancers per 100 people).

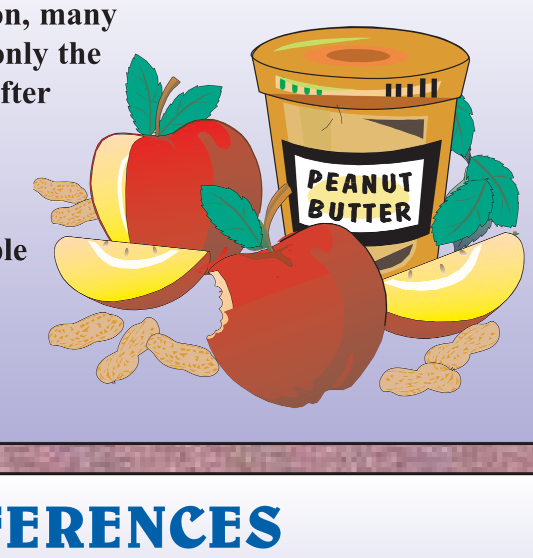


DISCUSSION AND CONCLUSIONS

There were not enough participants using each software package to fully evaluate the usefulness of each as policy tools. Is it worth doing such an evaluation with more subjects?

Based on their qualitative statements, *Crystal Ball* users generally seemed the most satisfied with the visual appearance of their output. (Figure 2) However, people drew conclusions from these figures without regard to such details as axis scaling which the software packages automatically set to incorporate all data on the plot unless the user overrides this option.

While users claimed they understood the purpose of MC simulation, many reported and examined only the means in their analysis after using software. Further research is needed to determine how to better ask questions to get people to respond to the full range of MC simulation results.



REFERENCES

Finkel, A.M. Toward less misleading comparisons of uncertain risks: the example of aflatoxin and Alar. *Environmental Health Perspectives,* 103(4): 376-385, 1995.

Kammen, D.M. and Hassenzahl, D.M. *Should We Risk It?: Exploring Environmental Health and Technical Problem Solving.* New Jersey: Princeton University Press, 1999.

FIGURE 3. Most meaningful and least meaningful output from *Modelmaker*. This software produced the greatest variation in user output.

A) One user (ID: MM01) who reported an ease of data manipulation score 5 produced this histogram plot. (Note: risk values are x1000 due to axis errors with low values).

B) The user that produced this histogram plot (ID: MM04) gave the software an ease of data manipulation score 2.